UNCLASSIFIED

AD 230 977

UNCLASSIFIED

STUDIES IN RESEARCH METHODOLOGY

I. COMPATIBILITY OF PSYCHOLOGICAL MEASUREMENTS
WITH PARAMETRIC ASSUMPTIONS

James V. Bradley

Aerospace Medical Laboratory

SEPTEMBER 1959

WRIGHT AIR DEVELOPMENT CENTER
AIR RESEARCH AND DEVELOPMENT COMMAND
UNITED STATES AIR FORCE
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

STUDIES IN RESEARCH METHODOLOGY

I. COMPATIBILITY OF PSYCHOLOGICAL MEASUREMENTS WITH PARAMETRIC ASSUMPTIONS

James V. Bradley

Aerospace Medical Laboratory

SEPTEMBER 1959

WRIGHT AIR DEVELOPMENT CENTER
AIR RESEARCH AND DEVELOPMENT COMMAND
UNITED STATES AIR FORCE
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

FOREWORD

ABSTRACT

The compatibility of typical psychological measurements with the assumptions of common, parametric, statistical tests is examined. Empirically obtained distributions of time scores and mathematically derived error distributions are used to illustrate conditions which give rise to serious violations of assumptions.

PUBLICATION REVIEW

This report has been reviewed and is approved.

FOR THE COMMANDER:

Walter F. Grether

WALTER F. GRETHER
Director of Operations
Aerospace Medical Laboratory

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

iv

INTRODUCTION

This is the first in a series of studies designed to give some concrete, quantitative indication of the degradation of statistical precision accompanying violation of parametric assumptions. By showing the effect of specific violations and the efficacy of certain "remedies" for them in certain completely defined cases, it is hoped to provide some of the perspective necessary for the proper selection and use of a statistical test, whether it be parametric or distribution-free.

The belief appears to be widespread that, while mild violations of parametric assumptions are common enough, extreme violations are quite rare, so much so, in fact, that one need hardly concern himself with their eventuality. There is little point in discussing the statistical effect of violations until the extent of their occurrence is appreciated. The first report in this series, therefore, will have the very limited objective of presenting "data", i.e. distributions, containing serious violations of assumptions and demonstrating how naturally and logically such violations can occur. Measurements typical of research in the area of experimental psychology will be used, namely time scores and errors. The distributions to be presented are distributions of scores for a single subject. Their relevance to multi-subject experiments will be treated in the Discussion section.

VIOLATIONS OF ASSUMPTIONS BY TIME SCORES

In a recent experiment, reported elsewhere (1) in detail, subjects were required to reach through a constant distance and operate the middle push button in a closely spaced array of three. Manipulated variables were the diameter and spacing of the push buttons and the orientation of the linear array; performance measures were reach-and-operation time, and "errors", i.e. frequency of inadvertent contact with the adjacent push buttons (which the subject was instructed to avoid touching).

1

In order to check the influence of experimental conditions upon the
distribution of performance measures (and therefore to check the validity
of the parametric assumptions of normality and homogeneity) extensive
distributions of time-scores were obtained, for a single subject, under
each of two of the conditions investigated in the experiment outlined
above. The subject was a full-time graduate assistant at an experimental
laboratory. He had had extensive experience in running subjects in
psychological experiments and had participated as subject in many such
experiments, some of which were conducted by the writer. His general
level and pattern of performance were therefore known; the level was, in
fact, high and the pattern, consistent. His motivation also was high;
the subject apparently regarded every experiment in which he participated
as a challenge. He was, in short, a "good" subject. Figure 1 shows a
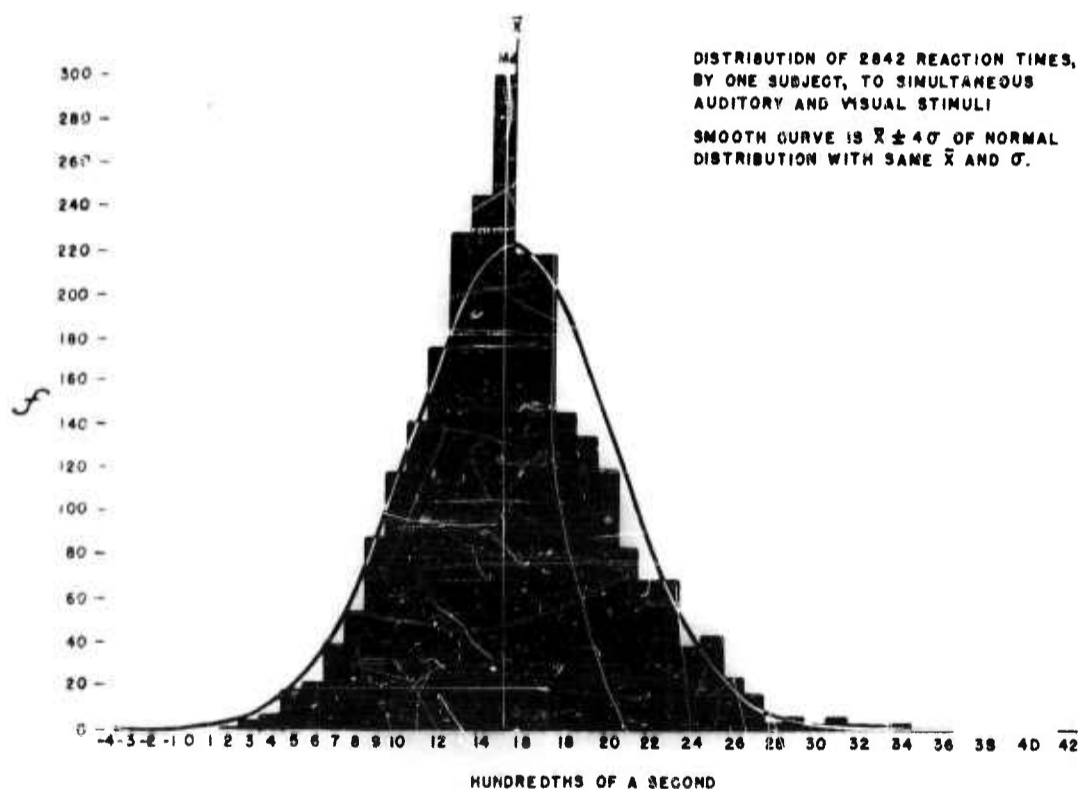
Figure 1. Nearly Normal Distribution of Reaction Times Generated
by same Subject Producing Distributions I and II (to follow).

2

very nearly normal distribution of time scores generated by this subject in a different and later experiment. It is presented as suggestive evidence that the nonnormalities of the distributions to be reported in the present experiment are "attributable" to the experimental conditions rather than to unique properties of the subject.

A single condition from the push button experiment was selected and the subject was given 2520 trials under that condition. These trials were administered in six experimental sessions of 420 trials each, each session being conducted on a different day and being interrupted by a five-minute break after the 210th trial. Subsequently, this entire procedure was repeated for a different experimental condition, using the same subject but a different experimenter. Evidences of sequential effects were checked by Cox and Stuart's S_2 sign test for trend (2) and, though small, were found to be highly statistically significant.

The first distribution obtained, Distribution I, is shown in Figure 2. The distribution was markedly bimodal with the proportion of time scores accompanied by errors being much greater for the second mode than for the first. It was hypothesized, therefore, that: (a) the first mode represented the short operation time made possible by a direct hit upon the push button "target" with the first thrust of the forefinger, the relatively few accompanying errors being caused by the "avoided" push buttons being contacted simultaneously by other portions of the subject's hand (or by the forefinger subsequent to operation of the center button); (b) the virtually scoreless trough between modes represented the time consumed in thrusting the finger at the target, missing it and perhaps striking the chassis or an avoided push button, then withdrawing and repoising the finger for a second thrust; (c) the second mode represented the time consumed in (b) plus the additional time required, after repoising the finger, to make a second, successful, thrust. (A miss on the first thrust did not necessarily result in an error, i.e., contact with an avoided button; and the subject could, and did occasionally, miss more than once during a trial).

3

FREQUENCY DISTRIBUTION, FOR ONE SUBJECT, OF TIME SCORES FOR THE CONDITION IN WHICH THE MIDDLE OF THREE 1/2" DIAMETER PUSH BUTTONS, SPACED 1/4" BETWEEN EDGES IN A VERTICAL ARRAY, IS REACHED TO AND OPERATED WHILE AVOIDING THE ADJACENT PUSH BUTTONS.

PORTION OF THE TOTAL FREQUENCY IN WHICH ONE OR BOTH OF THE ADJACENT PUSH BUTTONS WAS INADVERTENTLY OPERATED ▪ ▬▬

INADVERTENTLY TOUCHED, BUT NOT OPERATED ▪ ▭▭

Figure 2. Distribution I

In order to check this hypothesis a second distribution of 2520 scores, Distribution II, was obtained for the same subject in precisely the same way except that the diameter of the push buttons was increased from 1/2 inch to 1 inch and a different experimenter served as recorder. It was expected that increasing the size of the target would reduce the proportion of misses and therefore reduce the proportionate size of the second mode. In order to obtain exact rather than approximate

4

information as to which trials required more than one thrust, the subject reported this information and the experimenter recorded it for every trial. The second mode was vestigial, appearing as a very slightly humped, long positive tail, practically all of whose scores corresponded to trials in which the target was missed on the first thrust. (See Figure 3.) The hypothesis enunciated in the preceding paragraph was therefore considered as substantiated.
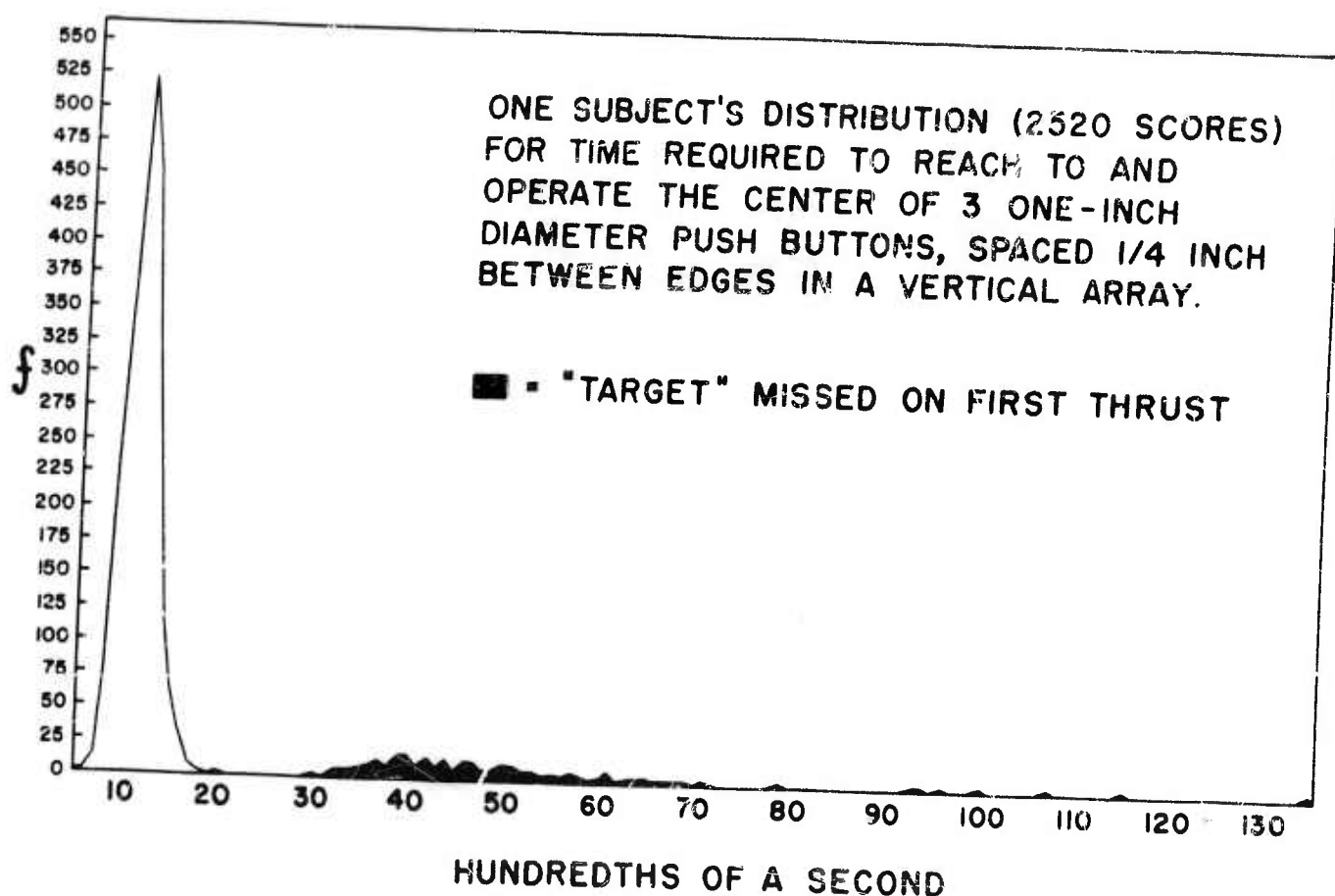
Figure 3. Distribution II

(Note change of scale: Maximum frequency is over twice that for Distribution I, and time-score range is smaller.)

Because of the large number of scores upon which they are based,
Distributions I and II may be regarded as reproducing the essential forms
of the time-score populations associated with the respective experimental
conditions under which they were obtained. Not only are they decidedly
nonnormal (see Figures 4 and 5), but their shapes, especially the second
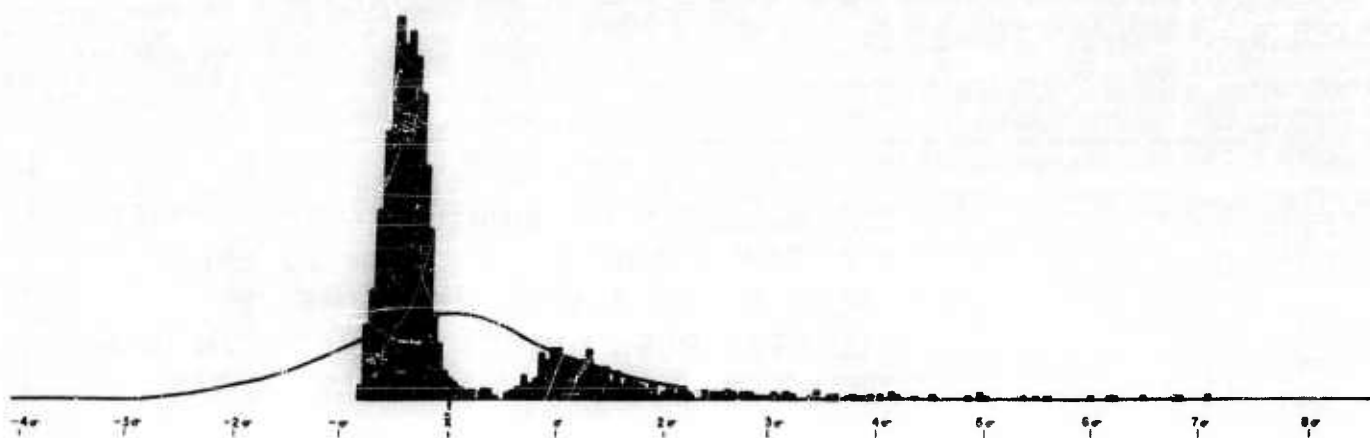
Figure 4. Distribution I (histogram) and Normal Distribution
with same Mean, Variance and Area

mode have proved to be a function of an experimental variable, namely
diameter, thus virtually assuring heterogeneity of variance. (Variances,
in fact, are 338.33 and 146.91 respectively and therefore are clearly
heterogeneous.) The assumptions common to most parametric statistical
tests are therefore violated by the "populations" of time scores
associated with the experiment. Distributions I and II therefore serve
to illustrate the hazard of assuming normality or homogeneity of variance
without empirical check, especially where time scores are involved. It
is particularly important to note, in this context, that Distributions I
and II were in no way "contrived", but were obtained under the conditions
of a perfectly routine experiment.

6

-4σ -3σ -2σ -σ X̄ σ 2σ 3σ 4σ 5σ 6σ 7σ 8σ 9σ 10σ
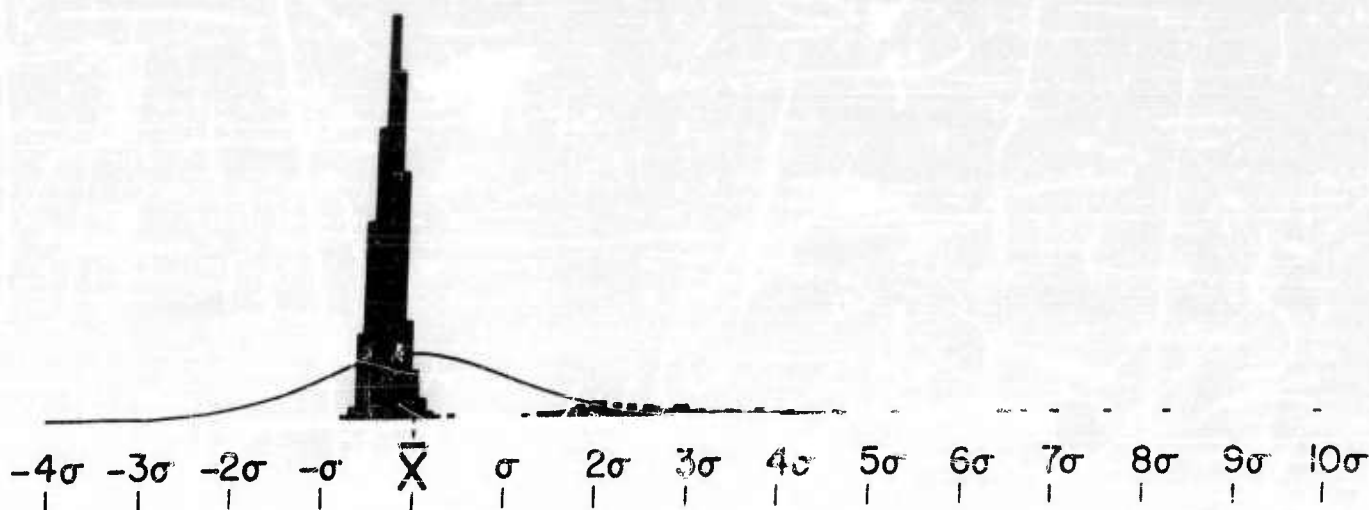
Figure 5. Distribution II (histogram) and Normal Distribution
with same Mean, Variance and Area

VIOLATION OF ASSUMPTIONS BY ERROR SCORES

The number, r, of errors committed in N trials is an error score. If
trials are randomly selected and their outcomes are independent, and if
the probability of an error on a single trial is a constant, p, then error
scores are binomially distributed with parameters N and p. The mean of
such a distribution is Np and its variance is $Np(1-p)$. The degree to
which such distributions violate the normality assumption is simply the
degree to which the normal approximation to the binomial distribution is
a poor fit, which, in turn, is a function of the parameters N and p. This
is illustrated in Figure 6.

As N decreases, the number of different values r can assume becomes
too small for the discrete binomial distribution to be well approximated
by the continuous normal distribution. The discrepancy is particularly
important at the tails of the distribution. In order to "fit" a normal
distribution with the same mean and variance, at its tails, the binomial

ERRORS ARE BINOMIALLY DISTRIBUTED.   THE NORMAL APPROXIMATION IS POOR IF THERE
IS AN APPRECIABLE PROBABILITY FOR EITHER ZERO ERRORS OR N ERRORS IN N TRIALS.
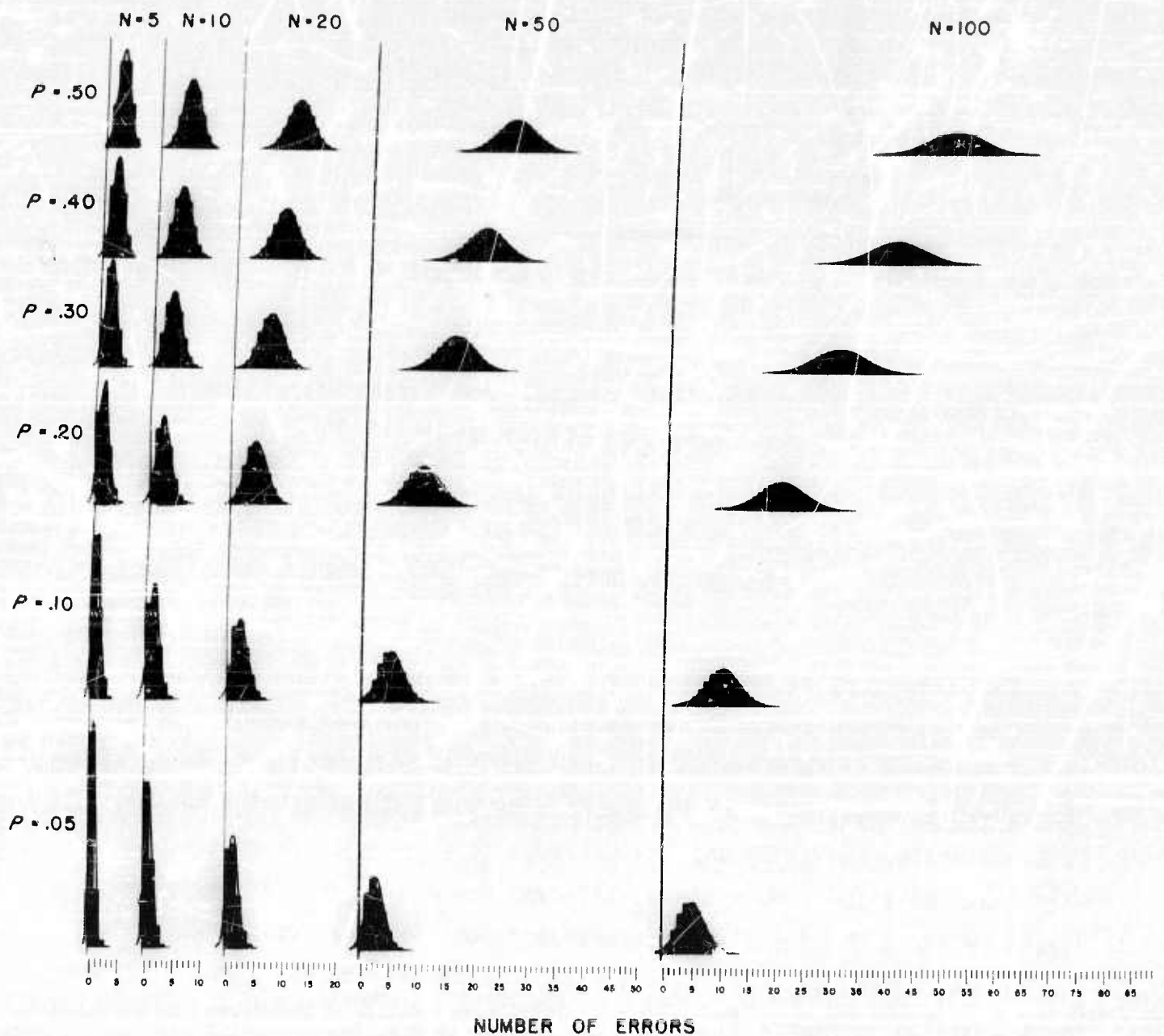
NUMBER OF ERRORS

DISTRIBUTIONS FOR NUMBER OF ERRORS IN N TRIALS WHEN P = PROBABILITY OF AN ERROR
ON A SINGLE TRIAL. SMOOTH CURVES ARE $\bar{X} \pm 3\sigma$ OF THE NORMAL APPROXIMATION WITH
DOTTED PORTION OF CURVE COVERING IMPOSSIBLE, $i.e.$ NEGATIVE, ERROR FREQUENCIES.

Figure 6. Error-Score Distributions

8

would have to approach zero probability by a series of fine gradations of diminishing discrete, i.e., point, probabilities. This it is unable to do if the number of point probabilities, i.e., bars in the histogram, is small.

As p departs increasingly from .5, the binomial distribution becomes increasingly asymmetrical. Thus, since the normal distribution is symmetrical, the assumption of normality is violated to increasing degree.

As N decreases and p approaches one of its extremes, zero or one, increasingly substantial proportions of the binomial histogram tend to become concentrated at that extreme. This forces more and more substantial proportions of the fitted, i.e., "assumed", normal distribution to be concentrated over error-score values which are, in fact, impossible, thus resulting in increasingly serious violations of the normality assumption.

It is clear, therefore, that error scores violate the parametric assumption of normality and that the degree of violation is likely to become appreciable if either the probability of an error on a single trial or the number of trials upon which the error score is based is small. If both are small the error distribution is certain to be quite appreciably nonnormal.

These conditions are, in fact, quite likely to obtain in practice. In the previously referenced multi-subject push-button experiment (1) the empirical probability of an error on a single trial, i.e., the obtained proportion of errors, ranged from .49 to .02 depending on the experimental condition. (For errors defined as inadvertent operation, rather than touching, of adjacent push buttons, empirical probabilities ranged from .06 to .00.) It is natural to define as errors events having low probability of occurrence, p. This, coupled with the widespread tendency to program an experiment so as to obtain from each subject a small number of trials, N, under each of a large number of treatments, tends to create the conditions under which the normality assumption is appreciably violated.

9

DISCUSSION

Neither time scores nor errors satisfy the parametric assumption of normality. A normal distribution has infinite range, is symmetrical, and is continuously distributed. None of these properties is characteristic of raw, absolute time scores or errors:

Range: The normal distribution extends from minus infinity to plus infinity. Absolute time scores, however, cannot be negative, and, in fact, generally cannot drop below some positive value corresponding to a physiological limit for the speed with which the task can be performed. Error scores, i.e., the number of errors in N trials, cannot be less than zero nor greater than N, (assuming a maximum of one error per trial). If an appreciable proportion of a normal curve "fitted" to a time score or error distribution covers values which are in fact impossible, then the normality assumption has been appreciably violated.

Symmetry: While it is not impossible for time score or error distributions to be exactly symmetrical, it is unlikely. Time scores tend to be positively skewed, presumably owing to the fact that there is no limit upon the value which can be assumed by scores above the median, while those below the median must be concentrated between the median and the physiological limit. Error scores are positively skewed if p is less than $1/2$ and negatively skewed if it exceeds $1/2$, the degree of skewness (for constant N) increasing with increasing values of $|p - 1/2|$.

Continuity: While elapsed time is continuously distributed, measured time has a discrete distribution owing to the fact that infinite precision of measurement is impossible. In the experiment reported herein, for example, the time clock was calibrated in hundredths of a second, thus giving a discrete distribution of measurements, time values between points, i.e., hundredths, being recorded as if "belonging" to the nearest point. (Interpolation would merely increase the fineness of the discrete gradations, e.g., substituting interpolated thousandths for recorded

10

hundredths.) Usually the measuring instrument is capable of sufficient gradation to render this criticism trivial; however, there are exceptions. In the case of errors, discontinuity may be a serious contributor to degree of nonnormality. The number of errors in N trials can assume only ·N + 1 different values. If N is small, the point probabilities for these N + 1 values must change by gross steps rather than by the succession of fine gradations which would be necessary to approximate well the normal curve. This is particularly important at the tails of the "fitted" normal distribution where the gross step is likely to be from a relatively large value to zero. This sudden descent cannot be matched by a corresponding drop in the ordinate of the fitted normal curve because the normal curve must approach zero probability asymptotically.

In many cases the central portion of a distribution is well approximated by a fitted normal curve, the fit becoming increasingly poor as the tail areas are approached. If the fit is "good" over say 90 to 95% of the area covered by the curve, the curves tend, very deceptively, to give the general appearance of a good overall fit, thus tempting the experimenter to make a false proclamation of "normality". The fit at the tails, however, is of critical importance and has not received the attention it deserves. Extreme, i.e., tail, values from the hypothesized distribution are those which contribute the most toward giving to a test statistic the extreme values which would place it in its rejection region; that is to say, the greater the number of sample observations whose values correspond to those in a single tail of the hypothesized distribution, the more likely is the test statistic to fall in its rejection region. The fit at the tails is therefore of critical importance to the commission of Type I, and, indirectly, of Type II errors. If the fit at the tails is "poor" the true probability of such errors may differ greatly from their nominal probabilities read from tables which were constructed under the assumption of normality. As has been shown, poorness of fit, i.e., extreme nonnormality, at the tails can result from the presence of impossible scores close to the bulk of the distribution, as well as from pronounced asymmetry or limitations on the number of different values a score can assume.

11

The preceding data and discussion have concerned distributions of scores for a single subject. They are relevant to multi-subject populations (defined here as distributions composed of an infinite number of individual, i.e., not mean, scores each of which was obtained from a different subject) to the degree that the individual subjects' distributions resemble each other in form and central tendency. Naturally a certain variability among individual subjects' distributions is to be expected in both these respects, and this variability may tend to make the multi-subject distribution more nearly normal than the distributions for the subjects as individuals. For example, suppose each subject's distribution were identical in form to Distribution I, but different in location, i.e., mean. If the means all fell within a range of a few hundredths of a second, the multi-subject distribution, although perhaps more nearly bell-shaped than the individual subjects' distributions, would still be bimodal with a fairly sharply defined trough between modes. However, if the subjects' true means were evenly distributed over a range of 25 hundredths of a second, the troughs for some subjects would correspond to modes for other subjects, with the result that the multi-subject distribution would tend to be unimodal and somewhat less skewed. Whether or not a better approximation to normality is obtained in proceeding from a "typical" one-subject distribution to a multi-subject distribution would appear to depend roughly upon the relative variance and shape of the distribution of individual subjects' true means with respect to the typical one-subject distribution of individual scores. If the shape of the former is more nearly normal or if its variance is much smaller than that of the latter, then, with infrequent exceptions, one would expect the multi-subject distribution to be more nearly normal than the typical one-subject distribution.

While these considerations should not be discounted, neither should they be weighted too heavily. All of the previous comments relative to range and continuity apply with equal force to multi-subject distributions. (Although the range of time scores is undoubtedly greater in the multi-subject distribution, the comments continue to apply if "physiological limit" is now understood to refer to the fastest possible time by any subject.) A greater degree of symmetry might frequently be expected in the central portion of a multi-subject distribution. However, the presence of impossible scores within a few standard deviations of the median and on one side of it, or unequally distant from it, will still tend to insure asymmetry at the tails.

12

The data presented in the body and appendix of this report illustrate violations of two other parametric assumptions, homogeneity of variance and independent observations, i.e. uncorrelated scores. Heterogeneity of variance exists among both the time-score and error-score distributions, and scores of both time-score distributions are sequentially correlated.

The assumption of homogeneity of variance, in tests for equality of means, is a particularly frustrating one. Variances may be unequal when the null hypothesis of equal means is true, but they are particularly likely to be so when the null hypothesis is false, due to the fact that in many cases means and variances are positively correlated. Heterogeneity of variance, therefore, tends to suggest that the null hypothesis is false. The experimenter, however, does not "know" whether or not means are unequal until he performs the statistical test, and this he cannot do so long as variances are heterogeneous. Techniques, of course, are available for coping with this situation; however, they are cumbersome at best. The likelihood of correlation between means and variances is intuitively obvious in the case of time scores: the longer the time required to perform a given type of task, the greater its variability would be expected to be. This, in fact, was the case for the time score distributions, I and II, presented earlier. In the case of binomially distributed errors, correlation between mean, Np, and variance, $Np(1-p)$, is inevitable if either N or p is held constant and is quite likely in all cases. (See Figure 6)

An assumption made by nearly all statistical tests, whether parametric or distribution-free, is that observations are independent, i.e., that the outcome, or score obtained, from one trial is not influenced by that of any preceding trial. When more than one score is obtained from a single subject, the assumption of independence generally implies that there must be no sequential effects, i.e., no learning, no fatigue and no motivational fluctuations such as would be caused by boredom. Experiences of the writer and his colleagues suggest that this assumption is never fully met when repetitive measurements, subject to such sequential effects, are taken upon a single subject. After thousands of "practice" trials the subject is still learning, and by the time he has received that much practice, boredom and fluctuations of attention are likely to have appreciable influence upon his performance. Lack of independence due to such sequential effects can, of course, be avoided by using only one score from each subject, different and

13

nonoverlapping groups of subjects being used under the various experimental conditions.

CONCLUSION

The parametric assumptions of normal distributions, equal variances and uncorrelated scores are particularly susceptible to violation by measurements typical of research in experimental psychology. The extent of the violation may be small, or, in the case of the last two assumptions, there may be no violation. However, drastic violations may occur quite naturally as a logical consequence of entirely realistic experimental conditions, and such extreme violations are not at all uncommon.

BIBLIOGRAPHY

1. Bradley, J. V. & Wallis, R. A. Spacing of On-Off Controls,
   I: Push Buttons. WADC Technical Report 58-2, ASTIA Document
   No. 142272, April 1958.

2. Cox, D. R. & Stuart, A. Some Quick Sign Tests for Trend in
   Location and Dispersion. Biometrika, 1955, 42, 80-95.

3. National Bureau of Standards, Tables of Normal Probability Functions.
   Applied Mathematics Series 23, Washington, D. C.: U. S. Government
   Printing Office, 1953.

4. National Bureau of Standards, Tables of the Binomial Probability
   Distribution. Applied Mathematics Series 6, Washington, D. C.:
   U. S. Government Printing Office, 1949.

5. Staff of the Computation Laboratory, Harvard University, Tables of
   the Cumulative Binomial Probability Distribution. Cambridge,
   Massachusetts: Harvard University Press, 1955.

APPENDIX A

STATISTICAL TRENDS IN THE GENERATION OF THE ORIGINAL POPULATIONS

Distribution	Statistic	Session 1	2	3	4	5	6	All Sessions
I	Mean	28.38	24.83	27.31	28.07	28.37	25.78	27.12
	Median	20.99	20.09	21.07	21.13	17.85	16.86	20.00
	Mode	20	19	21	19	17	15	17
	σ	17.44	14.62	16.75	17.74	22.03	20.50	18.39
	Range	123	130	144	108	175	146	175
	Lowest Score	15	12	14	12	12	12	12
	Highest Score	137	141	157	119	186	157	186
	Day Recorded	10/12/56	13/12/56	14/12/56	15/12/56	19/12/56	20/12/56	
II	Mean	14.38	15.55	15.00	15.35	14.97	15.39	15.11
	Median	12.01	12.35	11.40	10.17	10.63	11.82	11.48
	Mode	11	12	11	10	9	11,12	11
	σ	9.56	11.04	11.03	15.07	13.49	11.70	12.12
	Range	62	86	56	129	109	89	130
	Lowest Score	8	9	6	7	7	8	6
	Highest Score	69	94	61	135	115	96	135
	Day Recorded	6/8/57	7/8/57	8/8/57	9/8/57	10/8/57	11/8/57	

For both distributions sequential effects were tested by means of Cox and Stuart's S_2 sign test for trend, applied to the entire distribution. For Distribution I, $Pr < 10^{-15}$; for Distribution II, $Pr < 10^{-15}$.

FIT AT TAILS BETWEEN TIME OR ERROR SCORE DISTRIBUTIONS AND NORMAL DISTRIBUTION

Proportion of Distribution Lying:

Distribution		Below			Above	
	$\overline{X}-3.090\sigma$	$\overline{X}-2.326\sigma$	$\overline{X}-1.645\sigma$	$\overline{X}+1.645\sigma$	$\overline{X}+2.326\sigma$	$\overline{X}+3.090\sigma$
Normal	.0010	.0100	.0500	.0500	.0100	.0010
Time-Score Dist. I	0	0	0	.0591	.0333	.0197
Time-Score Dist. II	0	0	0	.0975	.0594	.0256

Binomial Error-Score Distributions

p	N	$\overline{X}-3.090\sigma$	$\overline{X}-2.326\sigma$	$\overline{X}-1.645\sigma$	$\overline{X}+1.645\sigma$	$\overline{X}+2.326\sigma$	$\overline{X}+3.090\sigma$
.01	5	0	0	0	.0490	.0490	.0490
"	10	0	0	0	.0956	.0956	.0043
"	20	0	0	0	.1821	.0169	.0169
"	50	0	0	0	.0894	.0138	.0138
"	100	0	0	0	.0794	.0184	.0034
.05	5	0	0	0	.0226	.0226	.0226
"	10	0	0	0	.0861	.0115	.0115
"	20	0	0	0	.0755	.0159	.0026
"	50	0	0	0	.0378	.0118	.0032
"	100	0	0	.0371	.0631	.0115	.0043
.10	5	0	0	0	.0815	.0086	.0086
"	10	0	0	0	.0702	.0128	.0128
"	20	0	0	0	.0432	.0113	.0024
"	50	0	.0052	.0338	.0579	.0245	.0032
"	100	.0000	.0078	.0576	.0726	.0206	.0020
.20	5	0	0	0	.0579	.0067	.0067
"	10	0	0	0	.0328	.0328	.0064
"	20	0	0	.0692	.0867	.0100	.0026
"	50	.0002	.0057	.0480	.0607	.0144	.0023
"	100	.0003	.0057	.0469	.0558	.0113	.0016
.30	5	0	0	0	.0308	.0308	.0024
"	10	0	0	.0283	.0474	.0106	.0016
"	20	0	.0076	.0355	.0480	.0171	.0013
"	50	.0002	.0073	.0402	.0478	.0123	.0009
"	100	.0004	.0089	.0479	.0531	.0125	.0011
.40	5	0	0	.0778	.0870	.0102	0
"	10	0	.0061	.0464	.0548	.0123	.0017
"	20	.0005	.0036	.0510	.0565	.0065	.0016
"	50	.0008	.0057	.0540	.0573	.0076	.0014
"	100	.0006	.0084	.0399	.0423	.0100	.0009
.50	5	0	0	.0313	.0313	0	0
"	10	.0010	.0107	.0547	.0547	.0107	.0010
"	20	.0013	.0059	.0577	.0577	.0059	.0013
"	50	.0013	.0077	.0595	.0595	.0077	.0013
"	100	.0009	.0105	.0443	.0443	.0105	.0009

APPENDIX C

FIT AT TAILS: AREAS FARTHER THAN Z STANDARD DEVIATIONS FROM THE MEAN, EXPRESSED AS PERCENTAGES OF CORRESPONDING AREAS OF NORMAL DISTRIBUTION

Distribution		$Z=-3.090$	$Z=-2.326$	$Z=-1.645$	$Z=+1.645$	$Z=+2.326$	$Z=+3.090$
Normal		100	100	100	100	100	100
Time-Score Dist. I		0	0	0	118	333	1966
Time-Score Dist. II		0	0	0	195	594	2563
Binomial Error-Score Distributions							
p	N						
.01	5	0	0	0	98	490	4901
"	10	0	0	0	191	956	427
"	20	0	0	0	364	169	1686
"	50	0	0	0	179	138	1382
"	100	0	0	0	159	184	343
.05	5	0	0	0	45	226	2259
"	10	0	0	0	172	115	1150
"	20	0	0	0	151	159	257
"	50	0	0	0	76	118	319
"	100	0	0	74	126	115	427
.10	5	0	0	0	163	86	856
"	10	0	0	0	140	128	1280
"	20	0	0	0	86	113	239
"	50	0	52	68	116	245	322
"	100	3	78	115	145	206	198
.20	5	0	0	0	116	67	672
"	10	0	0	0	66	328	637
"	20	0	0	138	173	100	259
"	50	19	57	96	121	144	251
"	100	28	57	24	112	113	155
.30	5	0	0	0	62	308	243
"	10	0	0	57	95	106	159
"	20	0	76	71	96	171	128
"	50	17	73	80	96	123	93
"	100	40	89	96	106	125	109
.40	5	0	0	156	174	102	0
"	10	0	51	93	110	123	168
"	20	52	56	102	113	65	161
"	50	76	57	108	115	76	137
"	100	56	84	80	85	100	88
.50	5	0	0	63	63	0	0
"	10	98	107	109	109	107	98
"	20	129	59	115	115	59	129
"	50	130	77	119	119	77	130
"	100	89	105	89	89	105	89

FREQUENCY DISTRIBUTIONS

Score	Frequency Distr. I	Frequency Distr. II
6		1
7		11
8		107
9		252
10		378
11		523
12	7	450
13	47	319
14	73	123
15	123	60
16	174	23
17	251	7
18	232	1
19	241	
20	222	1
21	200	
22	151	
23	111	
24	55	
25	36	
26	17	
27	12	
28	10	
29	6	
30	4	1
31	4	
32	3	4
33	5	5
34	5	5
35	1	6
36		8
37	3	12
38	4	9
39	6	17
40	15	17
41	9	9
42	12	15
43	28	8
44	24	13
45	32	5
46	32	12
47	24	12
48	22	6
49	23	7
50	24	11

Score	Frequency Distr. I	Frequency Distr. II
51	31	11
52	20	6
53	19	7
54	26	4
55	10	5
56	17	4
57	7	6
58	11	4
59	9	1
60	7	3
61	4	7
62	8	1
63	1	2
64	3	2
65	6	4
66	2	1
67	4	1
68	6	1
69	2	1
70	3	
71	4	2
72	3	1
73	3	
74	3	
75	4	
76	3	
77		
78	3	
79	3	1
80	1	
81	1	
82	1	
83	3	
84	1	
85	3	
86	2	
87		
88		
89	1	
90	4	
91		
92	2	
93	2	1
94		1
95	1	

Score	Frequency Distr. I	Frequency Distr. II
96	2	1
97	1	
98	1	
99	2	
100	1	1
101	2	
102		
103	3	
104	2	
105		
106		
107	1	1
108		
109		
110	2	
111		
112		
113		
114		
115		1
116	1	
117		
118	3	
119	1	
120		
121		
122		
123		
124		
125	1	
126		
127	1	
128		
129	1	
130		
131		
132		
133		
134		
135		1
136		
137	1	
138		
139		
140	1	